



Steroidogenesiscorrected.ST25.txt
SEQUENCE LISTING

<110> Stocco, Douglas M
Clark, Barbara J
<120> Compositions and Methods for Regulation of Steroidogenesis
<130> 18491.40
<140> US 09/612,894
<141> 2000-07-10
<160> 19
<170> PatentIn version 3.2
<210> 1
<211> 1466
<212> DNA
<213> Mus sp.

RECEIVED
OCT 01 2003
TECH CENTER 1600/2900

<400> 1
gtcgaccac gcgtccgctc aggaccttga aaggctcagg aagaacaacc cttgagcacc 60
tcagcactca gcatgttcct cgctacgttc aagctgtgtg ctggaagctc ctatagacat 120
atgcggaata tgaaaggatt aaggcaccaa gctgtgctgg ccattggcca agagctcaac 180
tgagagcac tgggggattc cagtcccggt tggatgggtc aagttcgacg tcggagctct 240
ctgcttggtt ctcaactgga agcaacactc tatagtgacc aggagctgtc ctacatccag 300
cagggagagg tggctatgca gaaggccttg ggcatactca acaaccagga aggctggaag 360
aaggaaagcc agcaggagaa cggggacgaa gtgctaagta agatggtgcc agatgtgggc 420
aagggtgttc gcttggagggt ggtggtagac cagcccatgg acagactcta tgaagaactt 480
gtggaccgca tggaggccat gggagagtgg aacccaaatg tcaaggagat caaggctcctg 540
cagaggattg gaaaagacac ggtcatcact catgagctgg ctgcggcggc agcaggcaac 600
ctggtggggc ctgagactt cgtgagcgtg cgctgtacca agcgcagagg ttccacctgt 660
gtgctggcag gcatggccac acattttggg gagatgccgg agcagagtgg tgtcatcaga 720
gctgaacacg gccccacctg catggtgctt catccactgg ctggaagtcc ctccaagact 780
aaactcactt ggctgctcag tattgacctg aaggggtggc tgccgaagac aatcatcaac 840
caggtcctat cgcagaccca gatagagttc gcccaaccacc tgcgcaagcg cctggaagcc 900
agccctgcct ctgaggccca gtgttaagga ctgtccacca cattgacctg caaatcattg 960
gaagctctca caggaagcct gcaagtctgt ccatcttcag ctaacagat cgggaggggt 1020
ggtagtcagg agacactagg actgactggg aaaatcagga tcagcaaat agaaatgagg 1080
cttagaataa aagttctcta gtgtctccca ctgcatagct gtgaaggcta agggataagt 1140
agctatgaaa cctttcatct aggttgtat atgctgacct aaaagacacc agcagctacg 1200
aacaggggat gctaaggatc gggaaactgt gtcttaccag ctccaaatgt cactacctga 1260

Steroidogenesiscorrected.ST25.txt

aggcagtgtg cacacaaagc aaggtcttgc ctaggaaact ctgtaaaagt tctcctctgt 1320
 aaaaggccag aacttgaatg aaactaccta caaagggcct ttccagagta ttccaacttt 1380
 tctctgagga gaaatgaaac catcattgtg ccgacttccc tactaatccc atgacaataa 1440
 agaacataca taaaaaaaaa aaaaaa 1466

<210> 2
 <211> 284
 <212> PRT
 <213> Mus sp.

<400> 2

Met Phe Leu Ala Thr Phe Lys Leu Cys Ala Gly Ser Ser Tyr Arg His
 1 5 10 15

Met Arg Asn Met Lys Gly Leu Arg His Gln Ala Val Leu Ala Ile Gly
 20 25 30

Gln Glu Leu Asn Trp Arg Ala Leu Gly Asp Ser Ser Pro Gly Trp Met
 35 40 45

Gly Gln Val Arg Arg Arg Ser Ser Leu Leu Gly Ser Gln Leu Glu Ala
 50 55 60

Thr Leu Tyr Ser Asp Gln Glu Leu Ser Tyr Ile Gln Gln Gly Glu Val
 65 70 75 80

Ala Met Gln Lys Ala Leu Gly Ile Leu Asn Asn Gln Glu Gly Trp Lys
 85 90 95

Lys Glu Ser Gln Gln Glu Asn Gly Asp Glu Val Leu Ser Lys Met Val
 100 105 110

Pro Asp Val Gly Lys Val Phe Arg Leu Glu Val Val Val Asp Gln Pro
 115 120 125

Met Asp Arg Leu Tyr Glu Glu Leu Val Asp Arg Met Glu Ala Met Gly
 130 135 140

Glu Trp Asn Pro Asn Val Lys Glu Ile Lys Val Leu Gln Arg Ile Gly
 145 150 155 160

Lys Asp Thr Val Ile Thr His Glu Leu Ala Ala Ala Ala Gly Asn
 165 170 175

Leu Val Gly Pro Arg Asp Phe Val Ser Val Arg Cys Thr Lys Arg Arg
 180 185 190

Steroidogenesiscorrected.ST25.txt

Gly Ser Thr Cys Val Leu Ala Gly Met Ala Thr His Phe Gly Glu Met
195 200 205

Pro Glu Gln Ser Gly Val Ile Arg Ala Glu His Gly Pro Thr Cys Met
210 215 220

Val Leu His Pro Leu Ala Gly Ser Pro Ser Lys Thr Lys Leu Thr Trp
225 230 235 240

Leu Leu Ser Ile Asp Leu Lys Gly Trp Leu Pro Lys Thr Ile Ile Asn
245 250 255

Gln Val Leu Ser Gln Thr Gln Ile Glu Phe Ala Asn His Leu Arg Lys
260 265 270

Arg Leu Glu Ala Ser Pro Ala Ser Glu Ala Gln Cys
275 280

<210> 3
<211> 14
<212> PRT
<213> Mus sp.

<400> 3

Ala Glu His Gly Pro Thr Cys Met Val Leu His Pro Leu Ala
1 5 10

<210> 4
<211> 12
<212> PRT
<213> Mus sp.

<400> 4

Ala Leu Gly Ile Leu Asn Asn Gln Glu Gly Trp Lys
1 5 10

<210> 5
<211> 19
<212> PRT
<213> Mus sp.

<400> 5

Gly Ser Thr Cys Val Leu Ala Gly Met Ala Thr His Phe Gly Glu Met
1 5 10 15

Pro Glu Gln

Steroidogenesiscorrected.ST25.txt

<210> 6
 <211> 6
 <212> PRT
 <213> Artificial

<220>
 <223> Synthesized peptide of the StAR protein
 <400> 6

Asn Gln Glu Gly Trp Lys
 1 5

<210> 7
 <211> 9
 <212> PRT
 <213> Artificial

<220>
 <223> Synthesized peptide of the StAR protein
 <400> 7

Ala Glu His Gly Pro Thr Cys Met Val
 1 5

<210> 8
 <211> 11
 <212> PRT
 <213> Mus sp.

<400> 8

Ile Leu Asn Asn Gln Glu Gly Trp Lys Lys Glu
 1 5 10

<210> 9
 <211> 25
 <212> DNA
 <213> Mus sp.

<220>
 <221> misc_feature
 <222> (3)..(3)
 <223> n = (a or c or g or t/u) or (unknown or other)

<220>
 <221> modified_base
 <222> (6)..(6)
 <223> r = a or g

<220>
 <221> modified_base
 <222> (9)..(9)
 <223> y = c or t/u

<220>
 <221> misc_feature
 <222> (12)..(12)

<223> n = (a or c or g or t/u) or (unknown or other)

<220>

<221> misc_feature

<222> (15)..(15)

<223> n = (a or c or g or t/u) or (unknown or other)

<220>

<221> misc_feature

<222> (18)..(18)

<223> n = (a or c or g or t/u) or (unknown or other)

<220>

<221> modified_base

<222> (21)..(21)

<223> y = c or t/u

<400> 9

gcngarcayg gncnacntg yatgg

25

<210> 10

<211> 25

<212> DNA

<213> Mus sp.

<220>

<221> modified_base

<222> (5)..(5)

<223> r = a or g

<220>

<221> misc_feature

<222> (8)..(8)

<223> n = (a or c or g or t/u) or (unknown or other)

<220>

<221> misc_feature

<222> (11)..(11)

<223> n = (a or c or g or t/u) or (unknown or other)

<220>

<221> misc_feature

<222> (14)..(14)

<223> n = (a or c or g or t/u) or (unknown or other)

<220>

<221> modified_base

<222> (17)..(17)

<223> r = a or g

<220>

<221> modified_base

<222> (20)..(20)

<223> y = c or t/u

<220>

<221> misc_feature

<222> (23)..(23)

<223> n = (a or c or g or t/u) or (unknown or other)

<400> 10

ccatrcangt nggncrctgy tcngc

25

<210> 11
 <211> 17
 <212> DNA
 <213> Mus sp.

<220>
 <221> modified_base
 <222> (3)..(3)
 <223> y = c or t/u

<220>
 <221> modified_base
 <222> (6)..(6)
 <223> r = a or g

<220>
 <221> modified_base
 <222> (9)..(9)
 <223> r = a or g

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> n = (a or c or g or t/u) or (unknown or other)

<400> 11
 aaycarcarg gntggaa

17

<210> 12
 <211> 17
 <212> DNA
 <213> Mus sp.

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> n = (a or c or g or t/u) or (unknown or other)

<220>
 <221> modified_base
 <222> (9)..(9)
 <223> y = c or t/u

<220>
 <221> modified_base
 <222> (12)..(12)
 <223> y = c or t/u

<220>
 <221> modified_base
 <222> (15)..(15)
 <223> r = a or g

<400> 12
 ttccancryt cytgrrt

17

Steroidogenesiscorrected.ST25.txt

<210> 13
 <211> 401
 <212> DNA
 <213> Mus sp.

<400> 13
 aaccaggaag gctggaagaa ggaaagccag caggagaacg gggacgaagt gctaagtaag 60
 atggtgccag atgtgggcaa ggtgtttcgc ttggagggtg tggttagacca gcccatggac 120
 agactctatg aagaacttgt ggaccgcatg gaggccatgg gagagtggaa cccaaatgtc 180
 aaggagatca aggtcctgca gaggattgga aaagacacgg tcatcactca tgagctggct 240
 gcggcggcag caggcaacct ggtggggcct cgagacttcg tgagcgtgcg ctgtaccaag 300
 cgcagagggtt ccacctgtgt gctggcaggc atggccacac attttgggga gatgccggag 360
 cagagtgggtg tcatcagagc tgaacacggc cccacctgca t 401

<210> 14
 <211> 1466
 <212> RNA
 <213> Mus sp.

<400> 14
 gucgaccac gcguccgcuc aggaccuuga aaggcucagg aagaacaacc cuugagcacc 60
 ucagcacuca gcauguuccu cgcucguuc aagcugugug cuggaagcuc cuauagacau 120
 augcggaaaua ugaaaggauu aaggcaccaa gcugugcugg ccauuggcca agagcucaac 180
 uggagagcac ugggggguuc cagucccggg uggauugguc aaguucgacg ucggagcucu 240
 cugcuugguu cucaacugga agcaacacuc uauagugacc aggagcuguc cuacaucag 300
 cagggagagg uggcuangca gaaggccuug ggcauacuca acaaccagga aggcuggaag 360
 aaggaaagcc agcaggagaa cggggacgaa gugcuaagua agauggugcc agauggggc 420
 aagguguuuc gcuuggaggu ggugguagac cagcccaugg acagacucua ugaagaacuu 480
 guggaccgca uggaggccau gggagagugg aacccaaaug ucaaggagau caagguccug 540
 cagaggauug gaaaagacac ggucaucacu caugagcugg cugcggcggc agcaggcaac 600
 cugguggggc cucgagacuu cgugagcgug cgcuguacca agcgagagg uuccaccugu 660
 gugcuggcag gcauggccac acuuuuuggg gagaugccgg agcagagugg ugucaucaga 720
 gcugaacacg gccccaccug cauggugcuu cauccacugg cuggaagucc cuccaagacu 780
 aaacucacuu ggcugcucag uauugaccug aagggguggc ugccgaagac aaucaucaac 840
 cagguccuau cgcagacca gauagaguuc gcccaaccac ugcgcaagcg ccuggaagcc 900
 agcccugccu cugaggccca guguaagga cuguccacca cauugaccug caaaucuuug 960
 gaagcucuca caggaagccu gcaagucugu ccaucuucag cuaacagcau cgggaggggu 1020
 gguagucagg agacacuagg acugacuggu aaaaucagga ucagcaaaaau agaaaugagg 1080

Steroidogenesiscorrected.ST25.txt

cuuagaauaa aaguucucua gugucuccca cugcauagcu gugaaggcua agggauaagu	1140
agcuaugaaa ccuuucaucu aggcuuuguau augcugaccu aaaagacacc agcagcuacg	1200
aacaggggau gcuagggauc gggaacuguu gucuuaccag cuccaaaugu aacuaccuga	1260
aggcagugug cacacaaagc aaggucuugc cuaggaaacu cuguaaaagu ucuccucugu	1320
aaaaggccag aacuugaauug aaacuaccua caaagggccu uuccagagua uuccaacuui	1380
ucucugagga gaaaugaaac caucauugug cgcacuuccc uacuaauccc augacaauaa	1440
agaacauaca uaaaaaaaaa aaaaaa	1466

<210> 15
 <211> 722
 <212> DNA
 <213> Homo sapiens

<400> 15	
gtggattaac caggttcggc ggcggagctc tctactcggg tctcggctgg aagagactct	60
ctacagtac caggagctgg cctatctcca gcagggggag gaggccatgc agaaggcctt	120
gggcatcctt agcaaccaag agggctggaa gaaggagagt cagcaggaca atggggacaa	180
agtgatgagt aaagtgggtc cagatgtggg caaggtgttc cggctggagg tcgtggtgga	240
ccagcccatg gagaggctct atgaagagct cgtggagcgc atggaagcaa tgggggagtg	300
gaaccccaat gtcaaggaga tcaaggctct gcagaagatc ggaaaagata cattcattac	360
tcacgagctg gctgccgagg cagcaggaaa cctggtgggg ccccgtagct ttgtgagcgt	420
gcgctgtgcc aagcgccgag gctccacctg tgtgctggct ggcatggaca cagacttcgg	480
gaacatgcct gagcagaagg gtgtcatcag ggcggagcac ggtcccactt gcatggtgct	540
tcacccgttg gctggaagtc cctctaagac caaacttacg tggctactca gcatcgacct	600
caaggggtgg ctgccaaga gcatcatcaa ccaggctctg tcccagacct aggtggattt	660
tgccaaccac ctgcgcaagc gcctggagtc ccaccctgcc tctgaagcca ggtgttgaag	720
ac	722

<210> 16
 <211> 134
 <212> DNA
 <213> Homo sapiens

<400> 16	
atgctgctag cgacattcaa gctgtgcgct gggagctcct acagacacat gcgcaacatg	60
aaggggctga ggcaacaggc tgtgatggcc atcagccagg agctgaaccg gagggccctg	120
gggggccccca cccc	134

<210> 17
 <211> 19

steroidogenesiscorrected.ST25.txt

<212> DNA
 <213> Homo sapiens
 <400> 17
 actggaagcc tgcaagtct

19

<210> 18
 <211> 285
 <212> PRT
 <213> Homo sapiens
 <400> 18

Met Leu Leu Ala Thr Phe Lys Leu Cys Ala Gly Ser Ser Tyr Arg His
 1 5 10 15

Met Arg Asn Met Lys Gly Leu Arg Gln Gln Ala Val Met Ala Ile Ser
 20 25 30

Gln Glu Leu Asn Arg Arg Ala Leu Gly Gly Pro Thr Pro Ser Thr Trp
 35 40 45

Ile Asn Gln Val Arg Arg Arg Ser Ser Leu Leu Gly Ser Arg Leu Glu
 50 55 60

Glu Thr Leu Tyr Ser Asp Gln Glu Leu Ala Tyr Leu Gln Gln Gly Glu
 65 70 75 80

Glu Ala Met Gln Lys Ala Leu Gly Ile Leu Ser Asn Gln Glu Gly Trp
 85 90 95

Lys Lys Glu Ser Gln Gln Asp Asn Gly Asp Lys Val Met Ser Lys Val
 100 105 110

Val Pro Asp Val Gly Lys Val Phe Arg Leu Glu Val Val Val Asp Gln
 115 120 125

Pro Met Glu Arg Leu Tyr Glu Glu Leu Val Glu Arg Met Glu Ala Met
 130 135 140

Gly Glu Trp Asn Pro Asn Val Lys Glu Ile Lys Val Leu Gln Lys Ile
 145 150 155 160

Gly Lys Asp Thr Phe Ile Thr His Glu Leu Ala Ala Glu Ala Ala Gly
 165 170 175

Asn Leu Val Gly Pro Arg Asp Phe Val Ser Val Arg Cys Ala Lys Arg
 180 185 190

Arg Gly Ser Thr Cys Val Leu Ala Gly Met Ala Thr Asp Phe Gly Asn

Steroidogenesiscorrected.ST25.txt
200 205

195

Met Pro Glu Gln Lys Gly Val Ile Arg Ala Glu His Gly Pro Thr Cys
210 215 220

Met Val Leu His Pro Leu Ala Gly Ser Pro Ser Lys Thr Lys Leu Thr
225 230 235 240

Trp Leu Leu Ser Ile Asp Leu Lys Gly Trp Leu Pro Lys Ser Ile Ile
245 250 255

Asn Gln Val Leu Ser Gln Thr Gln Val Asp Phe Ala Asn His Leu Arg
260 265 270

Lys Arg Leu Glu Ser His Pro Ala Ser Glu Ala Arg Cys
275 280 285

<210> 19
<211> 1641
<212> DNA
<213> Homo sapiens

<400> 19
agaacaccag gtccaggctg cagctgcggg actcagaggc gaacgttgag gggctcagga 60
aggacgaaga accacccttg agagaagagg cagcagcagc gcggcagcag cagcggcagc 120
gacccaccca ctgccacatt tgccaggaaa caatgctgct agcgacattc aagctgtgcg 180
ctgggagctc ctacagacac atgcgcaaca tgaaggggct gaggcaacag gctgtgatgg 240
ccatcagcca ggagctgaac cggagggccc tggggggccc cacccttagc acgtggatta 300
accaggttcg gcggcggagc tctctactcg gttctcggct ggaagagact ctctacagt 360
accaggagct ggcctatctc cagcaggggg aggaggccat gcagaaggcc ttgggcatcc 420
ttagcaacca agagggtg aagaaggaga gtcagcagga caatggggac aaagtgatga 480
gtaaagtggc cccagatgtg ggcaagggtg tccggctgga ggtcgtggtg gaccagccca 540
tgagagggct ctatgaagag ctctgtggagc gcatggaagc aatgggggag tggaacccca 600
atgtcaagga gatcaaggct ctgcagaaga tcggaaaaga tacattcatt actcacgagc 660
tggtctccga ggcagcagga aacctggtg ggccccgtga ctttgtgagc gtgcgctgtg 720
ccaagrgcgg aggtccacc tgtgtgctg ctggcatggc cacagacttc gggaaacatgc 780
ctgagcagaa ggggtgtcatc agggcggagc acggtccac ttgcatggtg cttcaccctg 840
tggttggaag tccctctaag accaaactta cgtggctact cagcatccac ctcaaggggt 900
ggctgccccaa gagcatcatc aaccaggctc tgtcccagac ccagggtgat ttgccaacc 960
acctgcgcaa gcgcctggag tcccaccctg cctctgaagc caggtgttga agaccagcct 1020

steroidogenesiscorrected.ST25.txt

gctgttccca actgtgccca gctgcactgg tacacacgct catcaggaga atccctactg	1080
gaagcctgca agtctaagat ctccatctgg tgacagtggg atgggtgggg ttcgtgttta	1140
gagtatgaca ctaggattca gattggtgaa agtttttagt accaagaaaa cagggatgag	1200
ctcttggtt aaaaggtaac ttcattcact gattagctat gacatgaggg ttcaggccccg	1260
ctaaaaataa ttgtaaaact ttttttctgg gcccttatgt acccacctaa aaccatcttt	1320
aaaatgctag tggctgatat ggggtgtggg gatgctaacc acagggcctg agaagtcttg	1380
ctttatgggc tcaagaatgc catgcgctgg cagtacatgt gcacaaagca gaatctcaga	1440
gggtctctg cagccctctg ctctccccg ccgctgcaca gcaacaccac agaacaagca	1500
gcacccca gtagggtgct tccagaaata tagtccaagc tttctctgtg gaaaaagaca	1560
aaactcatta gtagacatgt ttccctattg ctttcatagg caccagtcag aataaagaat	1620
cataattcac aaaaaaaaaa a	1641



SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Stocco, Douglas M.
Clark, Barbara J.

(ii) TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATION OF
STEROIDOGENESIS

(iii) NUMBER OF SEQUENCES: 19

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
(B) STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
(C) CITY: Austin
(D) STATE: TX
(E) COUNTRY: U.S.A.
(F) ZIP: 78701

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE:
(B) COMPUTER:
(C) OPERATING SYSTEM:
(D) SOFTWARE:

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US Unknown
(B) FILING DATE: Concurrently Herewith
(C) CLASSIFICATION: Unknown

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Mayfield, Denise L.
(B) REGISTRATION NUMBER: 33,732
(C) REFERENCE/DOCKET NUMBER: 043375.0006/DLM

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 512/499-6200

(B) TELEFAX: 512/499-6290

5 (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1466 base pairs

(B) TYPE: nucleic acid

10 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA"

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

	GTCGACCCAC GCGTCCGCTC AGGACCTTGA AAGGCTCAGG AAGAACAACC CTTGAGCACC	60
	TCAGCACTCA GCATGTTCCCT CGCTACGTTT AAGCTGTGTG CTGGAAGCTC CTATAGACAT	120
20	ATGCGGAATA TGAAAGGATT AAGGCACCAA GCTGTGCTGG CCATTGGCCA AGAGCTCAAC	180
	TGGAGAGCAC TGGGGGATTC CAGTCCCGGG TGGATGGGTC AAGTTCGACG TCGGAGCTCT	240
25	CTGCTTGTTT CTCAACTGGA AGCAACACTC TATAGTGACC AGGAGCTGTC CTACATCCAG	300
	CAGGGAGAGG TGGCTATGCA GAAGGCCTTG GGCATACTCA ACAACCAGGA AGGCTGGAAG	360
	AAGGAAAGCC AGCAGGAGAA CGGGGACGAA GTGCTAAGTA AGATGGTGCC AGATGTGGGC	420
30	AAGGTGTTTC GCTTGGAGGT GGTGGTAGAC CAGCCCATGG ACAGACTCTA TGAAGAACTT	480
	GTGGACCGCA TGGAGGCCAT GGGAGAGTGG AACCCAAATG TCAAGGAGAT CAAGGTCCTG	540
35	CAGAGGATTG GAAAAGACAC GGTCATCACT CATGAGCTGG CTGCGGCGGC AGCAGGCAAC	600

CTGGTGGGGC CTCGAGACTT CGTGAGCGTG CGCTGTACCA AGCGCAGAGG TTCCACCTGT 660

GTGCTGGCAG GCATGGCCAC ACATTTTGGG GAGATGCCGG AGCAGAGTGG TGTCATCAGA 720

5 GCTGAACACG GCCCCACCTG CATGGTGCTT CATCCACTGG CTGGAAGTCC CTCCAAGACT 780

AAACTCACTT GGCTGCTCAG TATTGACCTG AAGGGGTGGC TGCCGAAGAC AATCATCAAC 840

10 CAGGTCCTAT CGCAGACCCA GATAGAGTTC GCCAACCACC TGCGCAAGCG CCTGGAAGCC 900

AGCCCTGCCT CTGAGGCCCA GTGTTAAGGA CTGTCCACCA CATTGACCTG CAAATCATTG 960

GAAGCTCTCA CAGGAAGCCT GCAAGTCTGT CCATCTTCAG CTAACAGCAT CGGGAGGGGT 1020

15 GGTAGTCAGG AGACACTAGG ACTGACTGGT AAAATCAGGA TCAGCAAAAT AGAAATGAGG 1080

CTTAGAATAA AAGTTCTCTA GTGTCTCCCA CTGCATAGCT GTGAAGGCTA AGGGATAAGT 1140

AGCTATGAAA CCTTTTCATCT AGGCTTGTAT ATGCTGACCT AAAAGACACC AGCAGCTACG 1200

20 AACAGGGGAT GCTAAGGATC GGGAAGTGT GTCTTACCAG CTCCAAATGT CACTACCTGA 1260

AGGCAGTGTG CACACAAAGC AAGGTCTTGC CTAGGAAACT CTGTAAAAGT TCTCCTCTGT 1320

25 AAAAGGCCAG AACTTGAATG AACTACCTA CAAAGGGCCT TTCCAGAGTA TTCCAACTTT 1380

TCTCTGAGGA GAAATGAAAC CATCATTGTG CCGACTTCCC TACTAATCCC ATGACAATAA 1440

AGAACATACA TAAAAAAAAA AAAAAA 1466

30

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 284 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

10	Met	Phe	Leu	Ala	Thr	Phe	Lys	Leu	Cys	Ala	Gly	Ser	Ser	Tyr	Arg	His	
	1				5					10					15		
	Met	Arg	Asn	Met	Lys	Gly	Leu	Arg	His	Gln	Ala	Val	Leu	Ala	Ile	Gly	
				20				25					30				
15	Gln	Glu	Leu	Asn	Trp	Arg	Ala	Leu	Gly	Asp	Ser	Ser	Pro	Gly	Trp	Met	
			35					40					45				
	Gly	Gln	Val	Arg	Arg	Arg	Ser	Ser	Leu	Leu	Gly	Ser	Gln	Leu	Glu	Ala	
		50					55					60					
20	Thr	Leu	Tyr	Ser	Asp	Gln	Glu	Leu	Ser	Tyr	Ile	Gln	Gln	Gly	Glu	Val	
	65					70				75					80		
	Ala	Met	Gln	Lys	Ala	Leu	Gly	Ile	Leu	Asn	Asn	Gln	Glu	Gly	Trp	Lys	
25				85					90						95		
	Lys	Glu	Ser	Gln	Gln	Glu	Asn	Gly	Asp	Glu	Val	Leu	Ser	Lys	Met	Val	
				100				105						110			
30	Pro	Asp	Val	Gly	Lys	Val	Phe	Arg	Leu	Glu	Val	Val	Val	Asp	Gln	Pro	
			115					120						125			
	Met	Asp	Arg	Leu	Tyr	Glu	Glu	Leu	Val	Asp	Arg	Met	Glu	Ala	Met	Gly	
35			130					135					140				

	Glu Trp Asn Pro Asn Val Lys Glu Ile Lys Val Leu Gln Arg Ile Gly	
	145	150 155 160
5	Lys Asp Thr Val Ile Thr His Glu Leu Ala Ala Ala Ala Gly Asn	
	165	170 175
	Leu Val Gly Pro Arg Asp Phe Val Ser Val Arg Cys Thr Lys Arg Arg	
	180	185 190
10	Gly Ser Thr Cys Val Leu Ala Gly Met Ala Thr His Phe Gly Glu Met	
	195	200 205
	Pro Glu Gln Ser Gly Val Ile Arg Ala Glu His Gly Pro Thr Cys Met	
	210	215 220
15	Val Leu His Pro Leu Ala Gly Ser Pro Ser Lys Thr Lys Leu Thr Trp	
	225	230 235 240
	Leu Leu Ser Ile Asp Leu Lys Gly Trp Leu Pro Lys Thr Ile Ile Asn	
20	245	250 255
	Gln Val Leu Ser Gln Thr Gln Ile Glu Phe Ala Asn His Leu Arg Lys	
	260	265 270
25	Arg Leu Glu Ala Ser Pro Ala Ser Glu Ala Gln Cys	
	275	280

(2) INFORMATION FOR SEQ ID NO:3:

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 14 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

5 Ala Glu His Gly Pro Thr Cys Met Val Leu His Pro Leu Ala
 1 5 10

(2) INFORMATION FOR SEQ ID NO:4:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 12 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

20 Ala Leu Gly Ile Leu Asn Asn Gln Glu Gly Trp Lys
 1 5 10

(2) INFORMATION FOR SEQ ID NO:5:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

35

Gly Ser Thr Cys Val Leu Ala Gly Met Ala Thr His Phe Gly Glu Met

1 5 10 15

Pro Glu Gln

5

(2) INFORMATION FOR SEQ ID NO:6:

- 10 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

20 Asn Gln Glu Gly Trp Lys

1 5

(2) INFORMATION FOR SEQ ID NO:7:

25

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 9 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

35

Ala Glu His Gly Pro Thr Cys Met Val

1

5

5 (2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

10 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ile Leu Asn Asn Gln Glu Gly Trp Lys Lys Glu

1

5

10

20 (2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

25 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA"

30

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: one-of(3, 12, 15, 18)

(D) OTHER INFORMATION: /mod_base= OTHER

35 /note= "N = (A or C or G or T/U) or (unknown or other)"

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: 6

(D) OTHER INFORMATION: /mod_base= OTHER

5 /note= "R = A or G"

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: one-of(9, 21)

10 (D) OTHER INFORMATION: /mod_base= OTHER

/note= "Y = C or T/U"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

15

GCNGARCAYG GNCCNACNTG YATGG

25

(2) INFORMATION FOR SEQ ID NO:10:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA"

(ix) FEATURE:

30

(A) NAME/KEY: modified_base

(B) LOCATION: one-of(5, 17)

(D) OTHER INFORMATION: /mod_base= OTHER

/note= "R = A or G"

35

(ix) FEATURE:

(A) NAME/KEY: modified_base
(B) LOCATION: one-of(8, 11, 14, 23)
(D) OTHER INFORMATION: /mod_base= OTHER
/note= "N = (A or C or G or T/U) or (unknown or other)"

5

(ix) FEATURE:

(A) NAME/KEY: modified_base
(B) LOCATION: 20
(D) OTHER INFORMATION: /mod_base= OTHER

10

/note= "Y = C or T/U"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCATRCANGT NGGNC CRTGY TCNGC

25

15

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

20

(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA"

(ix) FEATURE:

(A) NAME/KEY: modified_base

30

(B) LOCATION: 3

(D) OTHER INFORMATION: /mod_base= OTHER

/note= "Y = C or T/U"

(ix) FEATURE:

35

(A) NAME/KEY: modified_base

(B) LOCATION: one-of(6, 9)

(D) OTHER INFORMATION: /mod_base= OTHER

/note= "R = A or G"

5 (ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: 12

(D) OTHER INFORMATION: /mod_base= OTHER

/note= "N = (A or C or G or T/U) or (unknown or other)"

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AAYCARCARG GNTGGAA

17

15

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 base pairs

20

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

25

(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTCCANCCYT CYTGRTT

17

30

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

35

(A) LENGTH: 401 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

10	AACCAGGAAG GCTGGAAGAA GGAAAGCCAG CAGGAGAACG GGGACGAAGT GCTAAGTAAG	60
	ATGGTGCCAG ATGTGGGCAA GGTGTTTCGC TTGGAGGTGG TGGTAGACCA GCCCATGGAC	120
	AGACTCTATG AAGAACTTGT GGACCGCATG GAGGCCATGG GAGAGTGGAA CCCAAATGTC	180
15	AAGGAGATCA AGGTCCTGCA GAGGATTGGA AAAGACACGG TCATCACTCA TGAGCTGGCT	240
	GCGGCGGCAG CAGGCAACCT GGTGGGGCCT CGAGACTTCG TGAGCGTGCG CTGTACCAAG	300
20	CGCAGAGGTT CCACCTGTGT GCTGGCAGGC ATGGCCACAC ATTTTGGGGA GATGCCGGAG	360
	CAGAGTGGTG TCATCAGAGC TGAACACGGC CCCACCTGCA T	401

25 (2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1466 base pairs

(B) TYPE: nucleic acid

30 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "RNA"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

	GUCGACCCAC GCGUCCGCUC AGGACCUUGA AAGGCUCAGG AAGAACAACC CUUGAGCACC	60
5	UCAGCACUCA GCAUGUUCU CGCUACGUUC AAGCUGUGUG CUGGAAGCUC CUAUAGACAU	120
	AUGCGGAAUA UGAAAGGAUU AAGGCACCAA GCUGUGCUGG CCAUUGGCCA AGAGCUCAAC	180
	UGGAGAGCAC UGGGGGAUUC CAGUCCCGGG UGGAUGGGUC AAGUUCGACG UCGGAGCUCU	240
10	CUGCUUGGUU CUCAACUGGA AGCAACACUC UAUAGUGACC AGGAGCUGUC CUACAUCAG	300
	CAGGGAGAGG UGGCUAUGCA GAAGGCCUUG GGCAUACUCA ACAACCAGGA AGGCUGGAAG	360
15	AAGGAAAGCC AGCAGGAGAA CGGGGACGAA GUGCUAAGUA AGAUGGUGCC AGAUGUGGGC	420
	AAGGUGUUUC GCUUGGAGGU GGUGGUAGAC CAGCCCAUGG ACAGACUCUA UGAAGAACUU	480
	GUGGACCGCA UGGAGGCCAU GGGAGAGUGG AACCCAAUG UCAAGGAGAU CAAGGUCCUG	540
20	CAGAGGAUUG GAAAAGACAC GGUCAUCACU CAUGAGCUGG CUGCGGCGGC AGCAGGCAAC	600
	CUGGUGGGGC CUCGAGACUU CGUGAGCGUG CGCUGUACCA AGCGCAGAGG UUCCACCUGU	660
25	GUGCUGGCAG GCAUGGCCAC ACAUUUUGGG GAGAUGCCGG AGCAGAGUGG UGUCAUCAGA	720
	GCUGAACACG GCCCCACCUG CAUGGUGCUU CAUCCACUGG CUGGAAGUCC CUCCAAGACU	780
	AAACUCACUU GGCUGCUCAG UAUUGACCUG AAGGGGUGGC UGCCGAAGAC AAUCAUCAAC	840
30	CAGGUCCUAU CGCAGACCCA GAUAGAGUUC GCCAACCACC UGCGCAAGCG CCUGGAAGCC	900
	AGCCUGCCU CUGAGGCCCA GUGUUAAGGA CUGUCCACCA CAUUGACCUG CAAAUCAUUG	960
35	GAAGCUCUCA CAGGAAGCCU GCAAGUCUGU CCAUCUUCAG CUAACAGCAU CGGGAGGGGU	1020

	GGUAGUCAGG AGACACUAGG ACUGACUGGU AAAAUCAGGA UCAGCAAAAU AGAAAUGAGG	1080
	CUUAGAAUAA AAGUUCUCUA GUGUCUCCCA CUGCAUAGCU GUGAAGGCUA AGGGAUAAGU	1140
5	AGCUAUGAAA CCUUUCAUCU AGGCUUGUAU AUGCUGACCU AAAAGACACC AGCAGCUACG	1200
	AACAGGGGAU GCUAAGGAUC GGGAACUGUU GUCUUACCAG CUCCAAAUGU CACUACCUGA	1260
10	AGGCAGUGUG CACACAAAGC AAGGUCTUGC CUAGGAAACU CUGUAAAAGU UCUCCUCUGU	1320
	AAAAGGCCAG AACUUGAUG AAACUACCUA CAAAGGGCCU UUC CAGAGUA UUCAACUUU	1380
	UCUCUGAGGA GAAAUAGAAC CAUCAUUGUG CCGACUCCCC UACUAAUCCC AUGACAAUAA	1440
15	AGAACAUACA UAAAAAAAAA AAAAAA	1466

(2) INFORMATION FOR SEQ ID NO:15:

20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH:
- (B) TYPE:
- (C) STRANDEDNESS:
- (D) TOPOLOGY:

25 (ii) MOLECULE TYPE:

(A) DESCRIPTION:

SEQUENCE DESCRIPTION: SEQ ID NO:15:

30

267	GTGGATTAAC CAGGTTTCGGC GGCGGAGCTC TCTACTCGGT TCTCGGCTGG AAGAGACTCT	326
327	CTACAGTGAC CAGGAGCTGG CCTATCTCCA GCAGGGGGAG GAGGCCATGC AGAAGGCCTT	386
387	GGGCATCCTT AGCAACCAAG AGGGCTGGAA GAAGGAGAGT CAGCAGGACA ATGGGGACAA	446
447	AGTGATGAGT AAAGTGGTCC CAGATGTGGG CAAGGTGTTC CGGCTGGAGG TCGTGGTGGA	506
35	507 CCAGCCCATG GAGAGGCTCT ATGAAGAGCT CGTGGAGCGC ATGGAAGCAA TGGGGGAGTG	566

567 GAACCCCAAT GTCAAGGAGA TCAAGGTCCT GCAGAAGATC GGAAAAGATA CATTATTAC 626
 627 TCACGAGCTG GCTGCCGAGG CAGCAGGAAA CCTGGTGGGG CCCCCTGACT TTGTGAGCGT 686
 687 GCGCTGTGCC AAGCGCCGAG GCTCCACCTG TGTGCTGGCT GGCATGGACA CAGACTTCGG 746
 747 GAACATGCCT GAGCAGAAGG GTGTCATCAG GGCGGAGCAC GGTCCCCTT GCATGGTGCT 806
 5 807 TCACCCGTTG GCTGGAAGTC CCTCTAAGAC CAACTTACG TGGCTACTCA GCATCGACCT 866
 867 CAAGGGGTGG CTGCCAAGA GCATCATCAA CCAGGTCCTG TCCCAGACCC AGGTGGATTT 926
 927 TGCCAACCAC CTGCGCAAGC GCCTGGAGTC CCACCTGCC TCTGAAGCCA GGTGTTGAAG 986
 987 AC 988

10

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH:
- (B) TYPE:
- (C) STRANDEDNESS:
- (D) TOPOLOGY:

(ii) MOLECULE TYPE:

20

- (A) DESCRIPTION:

SEQUENCE DESCRIPTION: SEQ ID NO:16:

25

127 ATGCTGCTAG CGACATTCAA GCTGTGCGCT GGGAGCTCCT ACAGACACAT GCGCAACATG 186
 187 AAGGGGCTGA GGCAACAGGC TGTGATGGCC ATCAGCCAGG AGCTGAACCG GAGGGCCCTG 246
 247 GGGGGCCCCA CCCC 260

30

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH:
- (B) TYPE:
- (C) STRANDEDNESS:
- (D) TOPOLOGY:

(ii) MOLECULE TYPE:

(A) DESCRIPTION:

SEQUENCE DESCRIPTION: SEQ ID NO:17:

5

1051 ACTGGAAGCC TGCAAGTCT 1069

(2) INFORMATION FOR SEQ ID NO:18:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:

(B) TYPE:

(C) STRANDEDNESS:

15

(D) TOPOLOGY:

(ii) MOLECULE TYPE:

(A) DESCRIPTION:

20

SEQUENCE DESCRIPTION: SEQ ID NO:18:

1	MLLATFKLCA GSSYRHRNM KGLRQQAVMA ISQELNRRAL GGTPSTWIN	50
51	QVRRRSSLLG SRLEETLYSD QELAYLQQGE EAMQKALGIL SNQEGWKES	100
101	QQDNGDKVMS KVPDVGVKF RLEVVDQPM ERLYEELVER MEAMGEWNP	150
25	151 VKEIKVLQKI GKDTFITHL AAEAAGNLVG PRDFVSVRCA KRRGSTCVLA	200
	201 GMATDFGNMP EQKGVIRAEH GPTCMVLHPL AGSPSKTKLT WLLSIDLKGW	250
	251 LPKSIINQVL SQTQVDFANHL RKRLESHPA SEARC	285

(2) INFORMATION FOR SEQ ID NO:19:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:

(B) TYPE:

(C) STRANDEDNESS:

35

(D) TOPOLOGY:

(ii) MOLECULE TYPE:

(A) DESCRIPTION:

SEQUENCE DESCRIPTION: SEQ ID NO:19:

5

agaacaccag gtccaggctg cagctgcggg actcagaggc gaacgttgag gggctcagga 60
aggacgaaga accacccttg agagaagagg cagcagcagc gcggcagcag cagcggcagc 120
gacccccacca ctgccacatt tgccaggaaa caATGCTGCT AGCGACATTC AAGCTGTGCG 180
CTGGGAGCTC CTACAGACAC ATGCGCAACA TGAAGGGGCT GAGGCAACAG GCTGTGATGG 240
10 CCATCAGCCA GGAGCTGAAC CGGAGGGCCC TGGGGGGCCC CACCCCTAGC ACGTGGATTA 300
ACCAGGTTTC GCGGCGGAGC TCTCTACTCG GTTCTCGGCT GGAAGAGACT CTCTACAGTG 360
ACCAGGAGCT GGCCTATCTC CAGCAGGGGG AGGAGGCCAT GCAGAAAGGCC TTGGGCATCC 420
TTAGCAACCA AGAGGGCTGG AAGAAGGAGA GTCAGCAGGA CAATGGGGAC AAAGTGATGA 480
GTAAAGTGGT CCCAGATGTG GGCAAGGTGT TCCGGCTGGA GGTCGTGGTG GACCAGCCCA 540
15 TGGAGAGGCT CTATGAAGAG CTCGTGGAGC GCATGGAAGC AATGGGGGAG TGGAACCCCA 600
ATGTCAAGGA GATCAAGGTC CTGCAGAAGA TCGGAAAAGA TACATTCATT ACTCACGAGC 660
TGGCTGCCGA GGCAGCAGGA AACCTGGTGG GGCCCCGTGA CTTTGTGAGC GTGCGCTGTG 720
CCAAGCGCCG AGGCTCCACC TGTGTGCTGG CTGGCATGGC CACAGACTTC GGGAACATGC 780
CTGAGCAGAA GGGTGTGATC AGGGCGGAGC ACGGTCCAC TTGCATGGTG CTTACCCCGT 840
20 TGGCTGGAAG TCCCTCTAAG ACCAACTTA CGTGGCTACT CAGCATCGAC CTCAAGGGGT 900
GGCTGCCCAA GAGCATCATC AACCAGGTCC TGTCACAGAC CCAGGTGGAT TTTGCCAACC 960
ACCTGCGCAA GCGCCTGGAG TCCCACCCTG CCTCTGAAGC CAGGTGTTGA agaccagcct 1020
gctgttccca actgtgcccc gctgcactgg tacacacgct catcaggaga atccctactg 1080
gaagcctgca agtctaagat ctccatctgg tgacagtggg atgggtgggg ttcgtgttta 1140
25 gagtatgaca ctaggattca gattggtgaa agtttttagt accaagaaaa cagggatgag 1200
ctcttggatt aaaaggtaac ttcattcact gattagctat gacatgaggg ttcaggcccc 1260
ctaaaaataa ttgtaaaact tttttctggt gcccttatgt acccacctaa aaccatcttt 1320
aaaatgctag tggctgatat ggggtgtggg gatgctaacc acagggcctg agaagtcttg 1380
ctttatgggc tcaagaatgc catgcgctgg cagtacatgt gcacaaagca gaatctcaga 1440
30 ggggtctctg cagccctctg ctctcccg ccgctgcaca gcaacaccac agaacaagca 1500
gcacccca gtagacatgt tccagaaata tagtccaagc tttctctgtg gaaaaagaca 1560
aaactcatta gtagacatgt ttcctattg ctttcatagg caccagtcag aataaagaat 1620
cataattcac acaaaaaaaaa a 1641